Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 181 (2006) 21–28
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Alanine check points in HNN and HN(C)N spectra

Amarnath Chatterjee, Ashutosh Kumar, Ramakrishna V. Hosur *

Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India

Received 23 January 2006; revised 8 March 2006
Available online 30 March 2006

Abstract

Rapid resonance assignment is a key requirement in structural genomics research by NMR. In this context we present here two new pulse sequences, namely, HNN-A and HN(C)N-A that have been developed by simple modification of the previously described pulse sequences, HNN and HN(C)N [S.C. Panchal, N.S. Bhavesh, R.V. Hosur, Improved 3D triple resonance experiments, HNN and HN(C)N, for H^N and ^{15}N sequential correlations in (^{13}C, ^{15}N) labeled proteins: application to unfolded proteins, J. Biomol. NMR, 20 (2001) 135–147]. These increase the number of start/check points in HNN and/or HN(C)N spectra and hence help in pacing up resonance assignment in proteins.

Keywords: HNN-A; HN(C)N-A; Check points; Triplets of residues; Band selective pulse

1. Introduction

We recently described two three-dimensional triple-resonance experiments, HNN and HN(C)N, for rapid resonance assignment of H^N and ^{15}N nuclei along the backbone in (^{15}N, ^{13}C) doubly labeled proteins [1]. These were derived by some modification of HN(CA)NNH [2] and HN(COCA)NH [3–5] described previously. The salient features of the experiments are: (a) the three frequency axes are H^N (F_3), ^{15}N (F_1), and ^{15}N (F_2), (b) the experiments show correlations between amide and ^{15}N of three consecutive residues, $i, i-1$, and $i+1$ along the polypeptide chain, (c) the various peaks in the (F_1, F_3) and (F_2, F_3) planes of the 3D spectrum display different patterns of positive and negative signs depending upon whether or not any of the residues at i or $i-1$ position is a glycine, and (d) prolines are characteristically identifiable by the absence of one correlation peak. Glycines and prolines thus generate many check points and start points for sequential walk along the polypeptide chain. Using these ideas new protocols were developed for rapid sequential walk along the sequence [6,7]. Thus because of the many check points one can generally assign the backbone H^N and ^{15}N resonances without having to obtain the residue specific side chain assignments explicitly. As a result, during the last two years these sequences have been successfully applied to many proteins in the MW range 10–20 kDa [6–15]. These have included both folded proteins as well as unfolded proteins, though it was originally thought that the pulse sequences would be too insensitive for folded proteins.

Continuing such efforts for rapid resonance assignments, we have developed here a new set of pulse sequences, HNN-A and HN(C)N-A which helps generate new check points to aid sequential walks. These lead to fixed points based on the position of alanines along the sequence. Moreover as the glycine and proline check/start points remain unaltered the number of triplet patterns dramatically increases. These features thus strengthen the protocols based on HNN and HN(C)N [6,7], for resonance assignment and the backbone amide and ^{15}N assignments can be obtained much faster. These become extremely invaluable, when the polypeptide chain has far removed glycines along the sequence, which could pose some difficulties in the HNN based application alone.

We have experimentally tested these sequences with a few proteins, and we show here the data on DLC8, a component of the motor protein, dynein. We demonstrate the application of these to both folded and unfolded DLC8.

* Corresponding author. Fax: +91 22 2280 4610.
E-mail address: hosur@tifr.res.in (R.V. Hosur).

1090-7807/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2006.03.009

2. Materials and methods

2.1. Protein preparation

For preparation of the monomeric DLC8 protein, the plasmids containing the insert for the protein DLC8 were transformed into BL21-DE3 strain of *Escherichia coli*, the cells were grown at 37 °C in M9 minimal medium to A_{600} of ~0.6, and then induced for production of DLC8 using 0.4 mM IPTG. Uniformly ^{15}N and $^{15}N/^{13}C$ labeled protein samples were prepared by growing bacteria in M9 minimal media supplemented with 1 g/l^{-1} $^{15}NH_4Cl$ and 4 g/l^{-1} [$^{13}C_6$]glucose as the sole source of nitrogen and carbon, respectively, as required, in the solution. The cells were harvested by centrifugation at 4500 rpm for 20 min using Beckman Sorvall (GSA rotor) centrifuge. The cell pellet was resuspended in extraction buffer (20 mM Tris–HCl, pH 8.0 containing 200 mM NaCl, and 10 mM imidazole) containing 2.5 μg/ml leupeptin, 1 μg/ml pepstatin, 0.1 mM phenylmethylsulfonyl fluoride (PMSF), 1% Triton X-100, and 100 μg/ml lysozyme. After sonication (five shots of 15 s, constant duty cycle 60 Hz power, five times) and centrifugation (35,000 rpm, 4 °C, 45 min), the crude extract was purified by affinity chromatography on a Ni^{2+} column by eluting with 250 mM imidazole. Imidazole was removed by dialysis against extraction buffer without imidazole in a 3 kDa membrane. The $(His)_6$ tag was removed by treatment with thrombin protease for 12 h at room temperature. The purity of the protein sample was confirmed by SDS–PAGE. The NMR sample was prepared by concentrating the protein to ~1 mM, in acetate buffer (20 mM, pH 3.0) containing 250 mM NaCl and 1 mM dithiothreitol (DTT). For experiments with unfolded DLC8, the sample was concentrated to ~1 mM and exchanged with acetate buffer (20 mM, pH 3.0) containing 250 mM NaCl, 1 mM DTT, and 6 M guanidine hydrochloride, by ultra-filtration.

2.2. NMR spectroscopy

All NMR experiments were performed at 27 °C on a Varian *Inova* 600 MHz NMR spectrometer equipped with pulse shaping and pulse field gradient capabilities. For the HNN-A spectrum the delays T_N, and T_C, were both set to 28 ms. Fourty complex points were used along t_1(^{15}N) and t_2(^{15}N) dimensions and 1024 complex points along t_3 (H^N). Sixteen scans were used for each FID. The HN(C)N-A spectrum was recorded with the same T_N and T_C parameters, same number of t_1, t_2, t_3 points, and the T_{CC} delay was set to 9 ms.

3. Results and Discussions

3.1. HNN-A and HN(C)N-A pulse sequences

The new pulse sequences (Figs. 1B and D) are a result of an extremely simple modification of the previously described pulse sequences [1] HNN and HN(C)N (Figs. 1A and C). As indicated in the figure, one of the pulses is replaced by a band selective pulse. However, this simple modification results in substantial spectral changes. To appreciate these changes, a brief description of the HNN and HN(C)N sequences is presented below.

The HNN and HN(C)N experiments employ the following magnetization transfer pathways.

$$\begin{aligned}&\mathrm{HNN}: \mathrm{H}_i^{\mathrm{N}} \rightarrow \mathrm{N}_i(t_1) \overset{2T_N}{\rightarrow} \mathrm{C}^{\alpha}_{i-1,i} \overset{2\tau}{\rightarrow} \mathrm{N}_{i-1,i,i+1}(t_2) \rightarrow \mathrm{H}_{i-1,i,i+1}(t_3)\\ &\mathrm{HN(C)N}: \mathrm{H}_i^{\mathrm{N}} \rightarrow \mathrm{N}_i(t_1) \overset{2T_N}{\rightarrow} \mathrm{CO}_{i-1} \overset{2T_{CC}}{\rightarrow} \mathrm{C}^{\alpha}_{i-1} \overset{2\tau}{\rightarrow} \mathrm{N}_{i-1,i}(t_2) \rightarrow \mathrm{H}_{i-1,i}(t_3)\end{aligned} \tag{1}$$

$2T_N$, 2τ, and $2T_{CC}$ are the time periods during which magnetization transfers take place; 2τ should be chosen between 22–30 ms to get best transfers and obtain the characteristic peak patterns in the spectra. The choice would be a reasonable compromise between transfer efficiency and transverse relaxation losses.

In a chain of four residues, $i-2$ to $i+1$, the intensities of the diagonal (I_i^{d}) and cross (I_{i-1}^{c}, I_{i+1}^{c}) peaks in the (F_2, F_3) planes of the HNN spectrum are given by:

$$I_i^{\mathrm{d}} = -(E_1^2 E_3 E_9 K_{i1}^{\mathrm{d}} + E_2^2 E_5 E_{10} K_{i2}^{\mathrm{d}}), \tag{2}$$

$$\begin{aligned}I_{i-1}^{\mathrm{c}} &= E_1 E_4 E_7 E_9 K_{i-1}^{\mathrm{c}},\\ I_{i+1}^{\mathrm{c}} &= E_2 E_6 E_8 E_{10} K_{i+1}^{\mathrm{c}},\end{aligned} \tag{3}$$

where,

$$\begin{aligned}E_1 &= \cos p_i T_N \sin q_{i-1} T_N,\\ E_2 &= \sin p_i T_N \cos q_{i-1} T_N,\\ E_3 &= \cos p_{i-1}\tau \cos q_{i-1}\tau,\\ E_4 &= \sin p_{i-1}\tau \sin q_{i-1}\tau,\\ E_5 &= \cos p_i \tau \cos q_i \tau,\\ E_6 &= \sin p_i \tau \sin q_i \tau,\\ E_7 &= \sin p_{i-1} T_N \cos q_{i-2} T_N,\\ E_8 &= \cos p_{i+1} T_N \sin q_i T_N,\\ E_9 &= \cos n_{i-1}\tau,\\ E_{10} &= \cos n_i \tau\end{aligned} \tag{4}$$

and

$$\begin{aligned}&p_i = 2\pi {}^1J(\mathrm{C}_i^{\alpha} - \mathrm{N}_i); \quad q_i = 2\pi {}^2J(\mathrm{C}_i^{\alpha} - \mathrm{N}_{i+1});\\ &n_i = 2\pi {}^1J(\mathrm{C}_i^{\alpha} - \mathrm{C}_i^{\beta}),\end{aligned} \tag{5}$$

$$\begin{aligned}K_{i1}^{\mathrm{d}} &= \exp(-4T_N R_{2i}^N - 2\tau R_{2,i-1}^{\alpha}),\\ K_{i2}^{\mathrm{d}} &= \exp(-4T_N R_{2i}^N - 2\tau R_{2i}^{\alpha}),\\ K_{i-1}^{\mathrm{c}} &= \exp(-2T_N(R_{2i}^N + R_{2,i-1}^N) - 2\tau R_{2,i-1}^{\alpha}),\\ K_{i+1}^{\mathrm{c}} &= \exp(-2T_N(R_{2i}^N + R_{2,i+1}^N) - 2\tau R_{2i}^{\alpha}).\end{aligned} \tag{6}$$

^1J's and ^2J's represent, respectively, the one bond and two bond N–C^{α} coupling constants, R_2's are the various transverse relaxation rates. The expressions in Eq. (1) reduce to the same as given in Panchal et al. [1] except for the relaxation factors, under the assumptions: $p_i = p_{i-1}$, $q_i = q_{i-1}$, and $n_i = n_{i-1}$.

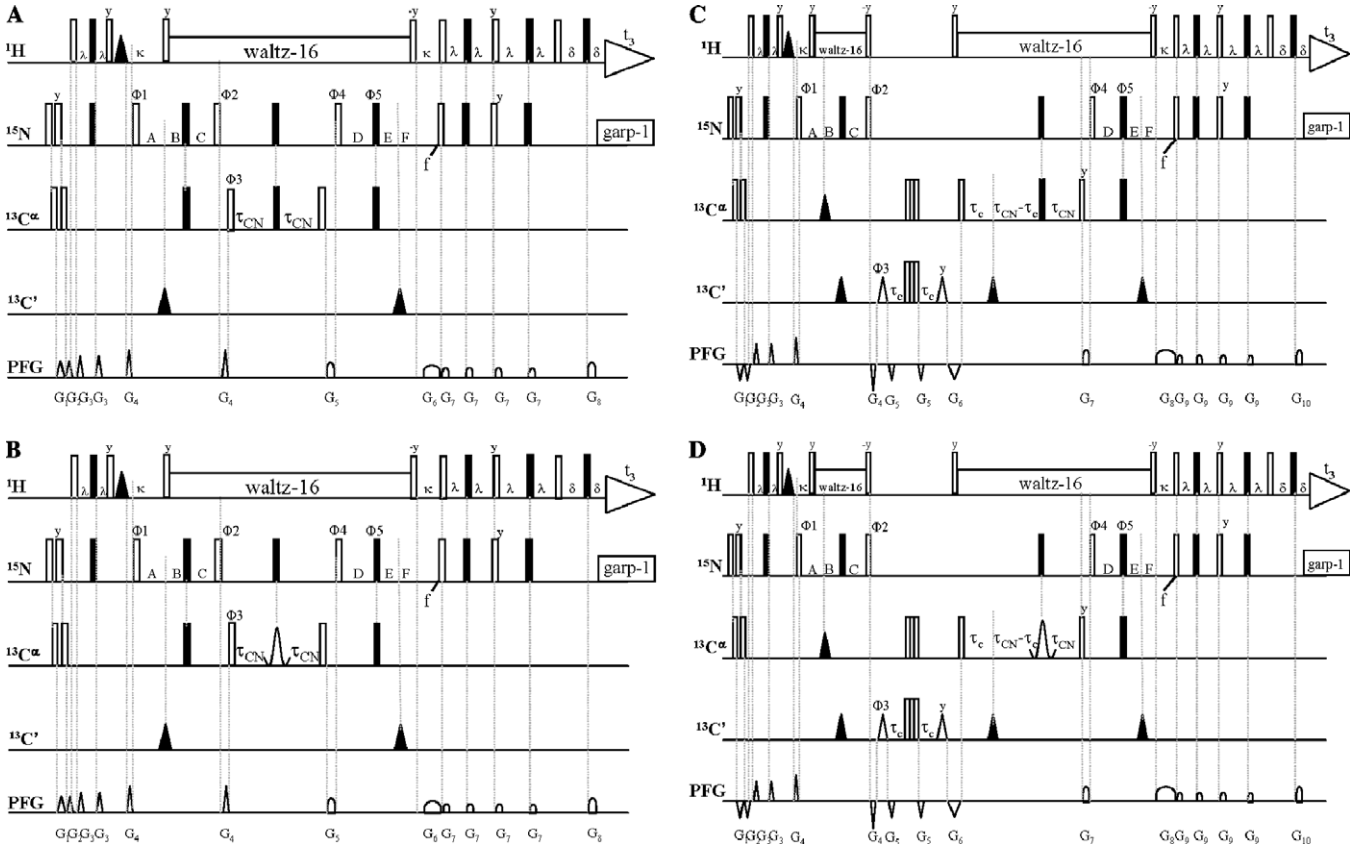

Fig. 1. Pulse sequence for the (A) HNN, (B) HNN-A, (C) HN(C)N, and (D) HN(C)N-A experiment. In all the cases, narrow (hollow) and wide (filled black) rectangular bars indicate non-selective 90° and 180° pulses, respectively. Unless indicated otherwise, the pulses are applied with phase x. The ^{1}H and ^{15}N carrier frequencies are set at 4.71 ppm (water) and 119.0 ppm, respectively. The ^{13}C carrier frequency is set at 56.0 ppm. The selective ^{1}H 90° pulse used for water flip-back during the first INEPT transfer is applied for a duration of 1.4 ms with one-lobe sinc profile. Proton decoupling using theWaltz-16 decoupling sequence with a field strength of 7.5 kHz is applied during most of the t_1 and t_2 evolution periods, and ^{15}N decoupling using the Garp-1 sequence with 0.5 kHz is applied during acquisition. The strength of the $^{13}C^{\alpha}$ pulses is adjusted so that they cause minimal excitation of carbonyl carbons. The band selective pulse in HNN-A (B) and HN(C)N-A (D) are indicated by bell shaped (hollow) pulses. The 180° ^{13}CO shaped pulse had a one-lobe sinc profile with minimal excitation of $^{13}C^{\alpha}$. The delays are $\lambda = 2.7$ ms, $\kappa = 5.4$ ms and $\delta = 0.35$ ms. τ_{CN} must be optimized and is around 12–16 ms. The values of the individual periods containing t_1 are: $A = t_1/2$, $B = T_N$, and $C = T_N - t_1/2$. The values of the individual periods containing t_2 are: $D = T_N - t_2/2$, $E = T_N$, and $F = t_2/2$. Phase cycling for the experiment is $\varphi_1 = 2(x)$, $2(-x)$; $\varphi_2 = x, -x, -x, x$; $\varphi_3 = \varphi_2$; $\varphi_4 = x$; $\varphi_5 = 4(x)$, $4(-x)$ and receiver $= 2(x)$, $2(-x)$. Frequency discrimination in t_1 is achieved using States-TPPI phase cycling of φ_1 along with the receiver phase; frequency discrimination in t_2 is achieved using the PEP sensitivity-enhanced gradient method. The N- and P-type signals are collected separately by inverting the sign of the G_6 gradient pulse (G_7 in case of HN(C)N). Signals recorded are manipulated post acquisition to generate pure absorption data. The gradient duration and levels are as follows: $G_1 = 0.5$ ms, 8 G/cm; $G_2 = 0.5$ ms, 5.6 G/cm; $G_3 = 0.5$ ms, 8 G/cm; $G_4 = 1$ ms, 11 G/cm; $G_5 = 1.0$ ms, 8 G/cm; $G_6 = 2.5$ ms, 26 G/cm; $G_7 = 0.5$ ms, 2 G/cm; $G_8 = 0.25$ ms, 25.8 G/cm. In (C) and (D) we have composite pulses, indicated by three continuous 90° pulses, [25] that can be used to evolve $^{13}C^{\alpha}$ and ^{13}CO under one-bond $^{13}C^{\alpha}$–^{13}CO coupling. The delay T_{CC} can be set to 4.5 ms. The values of the individual periods containing t_1 and t_2 periods are the same as in (A). Phase cycling for the experiment is $\varphi_1 = 2(-x)$, $2(x)$; $\varphi_2 = x, -x, -x, x$; $\varphi_3 = \varphi_2$; $\varphi_4 = x$; $\varphi_5 = 4(x)$, $4(-x)$; and receiver $= 2(x)$, $2(-x)$. The gradient duration and levels are as follows: $G_1 = 0.5$ ms, 8 G/cm; $G_2 = 0.5$ ms, 5.6 G/cm; $G_3 = 0.5$ ms, 8 G/cm; $G_4 = 1$ ms, 11 G/cm; $G_5 = 0.5$ ms, 7 G/cm; $G_6 = 1.0$ ms, 9 G/cm; $G_7 = 1.0$ ms, 8 G/cm; $G_8 = 2.5$ ms, 26 G/cm; $G_9 = 0.5$ ms, 2 G/cm; $G_{10} = 0.25$ ms, 25.8 G/cm.

Among the various evolutions and the transfers occurring through the pulse sequence, the evolution during the period 2τ is the most crucial from the sensitivity point of view, as, it is during this period alone that the magnetization resides on the C^{α} carbon which has the fastest transverse relaxation rate, and its minimization is crucial for highest sensitivity in the spectrum. In Fig. 2 we show the peak intensities under the above assumptions of equality of p's and q's, as a function of τ, for a set of p and q representing the β sheet structures; the average coupling constants for the α and β types of structures are slightly different [16]: for α helices, p and q are in the ranges 8–10 and 4–6 Hz, respectively, and for β structures these values are in the ranges 10–13 and 6–9 Hz, respectively.

C^{α} evolution during 2τ has another important consequence. A careful examination of the Eqs. (1)–(3) reveals that for the optimum choices of the transfer periods, the functions E_1–E_8 are all positive while, E_9, E_{10} are negative. This means depending upon presence or absence of E_9, E_{10} (which originate from evolution under C^{α}–C^{β} coupling during 2τ period) the cross peaks can have negative or positive signs, respectively. This is precisely the basis of generation of different patterns of peaks due to glycines which do not have a C^{β} carbon. These generate the check points

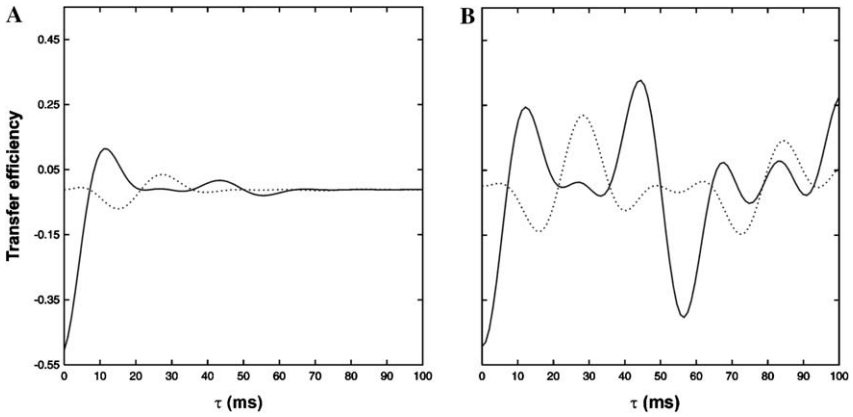

Fig. 2. Plots of the HNN coherence transfer efficiencies. The transfer functions for the diagonal peak, I^{d} (Eq. (2)) and the cross peak I^{c} (Eq. (3)) are described in the text. Here, (A) is for the transfer efficiencies calculated with relaxation terms while (B) is for calculations without the relaxation terms. The transfer efficiency is plotted as a function of τ. The plots were calculated by using, $J_{C^{\alpha}C^{\beta}}$, $J_{C^{\alpha}CO}$, J_{NCO} values of 35, 55, and 15 Hz, respectively. The $^{1}J_{C^{\alpha}N}$, $^{2}J_{C^{\alpha}N}$ values have been chosen to be 10.5 and 8.5 Hz, respectively. The value of T_N used in the transfer functions for HNN was 14.0 ms and T_{CC} was set to 4.5 ms. Thick and dotted lines represent diagonal and sequential peaks, respectively.

for sequential walks in the HNN spectrum. What this implies is that if a particular evolution under C^{α}–C^{β} coupling during 2τ period is prevented either by selective un-labeling of C^{β} carbon or by specific decoupling of C^{β}, then additional check points can be created in the 3D spectrum; the easy option here is to decouple the C^{β} and this is the basis of the present modification. These become particularly necessary if the protein sequence does not contain glycines for long stretches.

Among the various amino acid residues, the C^{β} chemical shifts of alanines are very distinct. The alanine C^{β}'s resonate most up-field (18–22 ppm) compared to all other residues (27–43 ppm). This is particularly more valid in unfolded proteins, where the chemical shift dispersions are rather small. Thus it is possible to apply band selective 180° pulses in the middle of the C^{α} evolution (2τ period in Figs. 1B and D) in the pulse sequences such that all C^{β} except those of alanines are inverted. Consequently, all the alanines will not have C^{α}–C^{β} coupling evolution and will behave like glycines; these lead to what one may call as 'alanine check points.' We hasten to add that these can be readily distinguished from the true glycine check points because of the distinctly different ^{15}N chemical shifts of alanines and glycines. This modified pulse sequence may be termed as HNN-A.

The expressions for the diagonal and the sequential peaks in HN(C)N at the same level of detail as in Eqs. (2)–(6) have already been described earlier [17] and we do not wish to repeat them here. The modification of HN(C)N to generate HN(C)N-A presented in Fig. 1D follow the same considerations as described for HNN-A.

Different kinds of band selective inversion pulses have been described in the literature [18–21]. For the alanine check points, an inversion band width of roughly 50 ppm (25–75 ppm) is required. Among the various possibilities, sech/tanh inversion pulses have the flattest inversion profiles over 90% of the band width [22]. However, these pulses are also reasonably long ranging from 1–5 ms. Square pulses with attenuated power can also be used. These pulses are short, simplest to calculate, but their excitation profiles are not uniformly flat. Nevertheless, we have used these pulses for the present demonstration, with a larger bandwidth specification so that the desired spectral region is properly inverted.

3.2. Peak patterns for proline, glycine, and alanine neighbors

As described earlier by Bhavesh et al. [6], in the present case also triplets of residues may be considered. In the HNN spectrum every F_1–F_3 and F_2–F_3 plane contains the diagonal peak ($F_1 = F_2 = N_i$) and two sequential peaks to $i-1$ and $i+1$ residues. On the other hand, in the HN(C)N spectrum, the F_2–F_3 plane contains the diagonal ($F_1 = F_2 = N_i$) peak and one sequential peak to $i-1$ residue, whereas the F_1–F_3 plane contains the diagonal peak and one sequential peak to $i+1$ residue. Thus although the HN(C)N sequence generates only i and $i-1$ correlations, the F_1–F_3 and F_2–F_3 planes taken together help identifying a triplet of consecutive residues.

We have already mentioned that the evolutions of the magnetization components are slightly different for glycine and non-glycine residues, because of the absence of the C^{β} carbon in the former. This results in different combinations of positive and negative signs for the various self and cross peaks in the different planes of the 3D spectra. Further the absence of an amide proton for a proline results in the absence of the corresponding peak. Moreover in the present case since the alanines also behave as glycines the list of peak patterns is expanded. In addition, ^{15}N chemical shifts also display certain residue type dependence [23]. Glycines are distinctly upfield while alanines are downfield compared to others. Interestingly, the average values of the shifts for the different residue types are similar in both folded and unfolded proteins, though the spreads are more in

the folded proteins. Thus, there will be different patterns of peaks for different triplets of residues containing glycines, alanines, and prolines. The expected peak patterns for each of the above cases in the F_1–F_3 planes of the HNN and HN(C)N spectra, at the F_2 chemical shift of the central residue are schematically shown in Fig. 3. In each of the planes, the peaks occur aligned at the amide (F_3) chemical shift of the central residue. The choice of the relative ^{15}N chemical shifts of Z, Z′, and X residues is quite arbitrary. In reality, the positions of the positive and negative peaks can get altered as per the relative chemical shifts. The important things to consider are: (i) the sign of the self or the diagonal ($F_1 = F_2$) peak and (ii) the signs of the sequential peaks relative to that of the diagonal peak.

Thus, here we have seven categories of triplets of residues that may be distinguished: (I) ZXZ′, (II) PXZ, ZXP, PXP, (III) PXA, PAX, PAA, AXP, AAP, XAP, PAP, (IV) XAZ, AXZ, ZXA, AAZ, ZAA, AXA, AAA, (V) GAP, PAG, PGA, AGP, (VI) GAX, AXG, XGA, XAG, GXA, AGX, and (VII) GGA, GAG, AGG, AAG, AGA, GAA, where X, Z, and Z′ can be any residue other than proline, glycine, and alanine. Category I is a general one, not containing glycines, alanines, and prolines, and has been included to be able to distinguish the special patterns from the general pattern, category II has prolines but no glycines, category III has combinations of alanines and prolines, category IV has alanines but no prolines and/or glycines, category V has combinations of glycines, alanines, and prolines, category VI has combination of glycines and alanines but no prolines, and category VII has combination of two glycines with one alanine and vice versa. Besides these the HNN-A and HN(C)N-A will also contain the earlier described [6] peak patterns due to combinations of glycines and prolines. The following salient features of the patterns would be helpful in proper analysis of the spectra.

In the HNN-A experiment which generates correlations from i to both $i-1$ and $i+1$ residues, the sign of the self peak for glycine and alanine is always opposite to that of any other residue; the actual signs will depend upon how the spectra are phased. The diagonal of glycine and alanines is chosen to have a negative sign. To emphasize this sign distinction, the diagonal peaks are shown with a different symbol (square) in the figure. This enables unambiguous discrimination between triplets having G or A, and triplets having X as the central residues. For example, the patterns for PAX, and PXA in HNN-A seem similar, but in the former the diagonal peak is negative and the sequential is positive, whereas the reverse is true in the latter case.

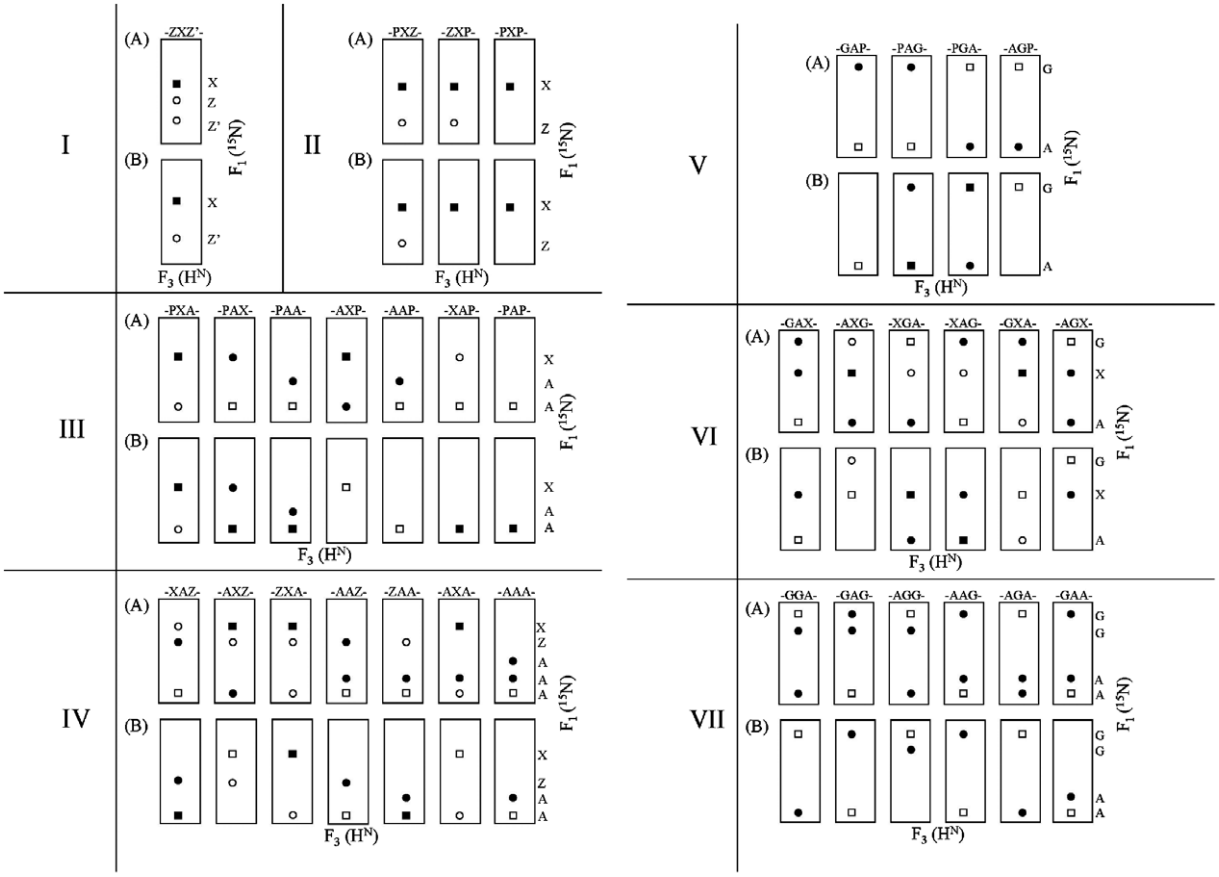

Fig. 3. Schematic patterns in the F_1–F_3 planes at the F_2 chemical shift of the central residue in the triplets mentioned on the top of each panel, in the HNN-A (A) and HN(C)N-A (B) spectra for various special triplet sequences of categories I–VII (see text). X, Z, Z′ is any residue other than glycine, alanine, and proline. Squares are the diagonal peaks and circles are the sequential peaks. Filled and open symbols represent positive and negative signals, respectively. In all cases the peaks are aligned at the F_3 (H^N) chemical shift of the central residue.

The sign of the sequential peak at $i-1$ position will be positive or negative depending upon whether that residue is glycine; alanine or otherwise. Similarly the sign of the sequential peak to $i+1$ residue will be positive or negative depending on whether the ith residue is glycine; alanine or otherwise.

In the HN(C)N-A experiment which generates i to $i-1$ correlation the signs of the self and sequential peaks are always opposite. The actual signs are dictated by whether the $i-1$ residue is a glycine or an alanine or otherwise, and of course by the phasing of the spectra. Again the diagonal has been chosen to be negative for glycine and alanine at $i-1$ position. Consequently, in the F_1–F_3 plane at F_2 chemical shift of residue i, the signs of the i and $i+1$ peaks are dictated by the nature of the residues at $i-1$ and i positions, respectively.

Comparison of the patterns in HNN-A and HN(C)N-A spectra enables ready discrimination of the $i-1$ and $i+1$ neighbors and this provides directionality to the assignment. The patterns in HN(C)N-A resolve some of the possible ambiguities in the HNN-A patterns and vice versa. For example, ZXP and PXZ can be readily distinguished

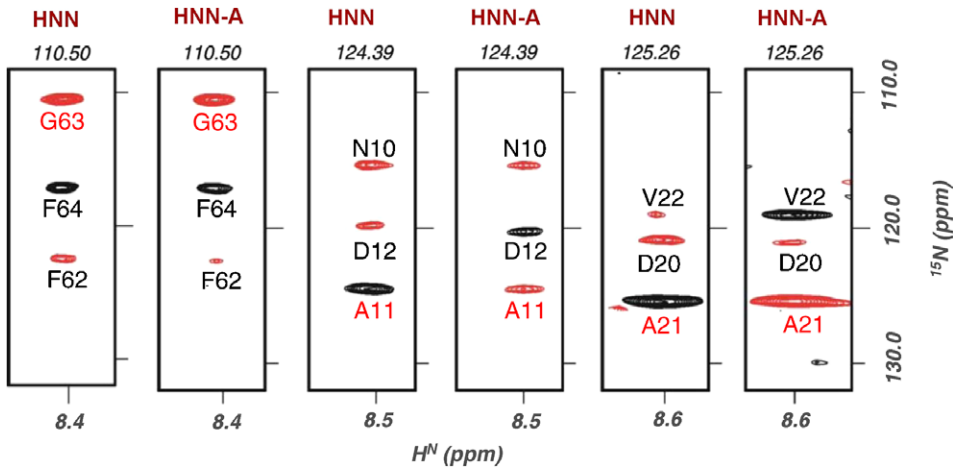

Fig. 4. Identical $F_1(^{15}N)$–$F_3(H^N)$ strips through HNN and HNN-A spectra of the protein DLC8 at pH 3, 27 °C. The numbers on top of each panel are the F_2 chemical shifts which help to identify the diagonal ($F_1 = F_2$) peaks. Black contours are positive peaks and red contours are negative peaks. The residues corresponding to the diagonal in each strip has been marked by red. The distinctive behavior of alanines is evident for the residues A11 and A21 in DLC8 as they show a negative diagonal in HNN-A.

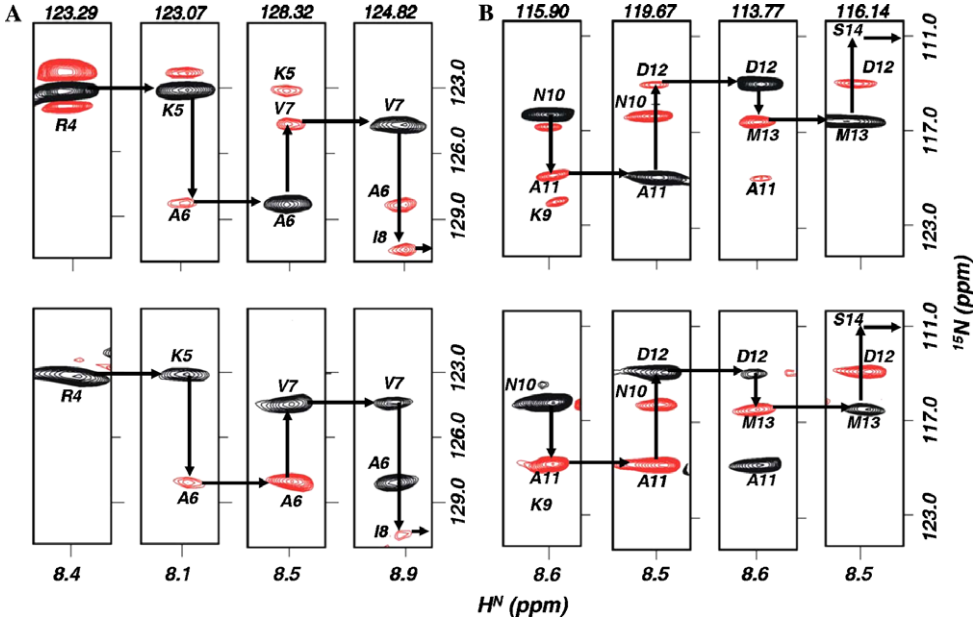

Fig. 5. Illustrative sequential walks through the HNN spectra (top panels) and the HNN-A spectra (bottom panel) for folded DLC8 (A) and unfolded DLC8 in 6 M guanidine hydrochloride (B). Black and red contours are positive and negative peaks, respectively. The strips, KAV, AVI in the bottom panel of (A) and NAD, ADM in the bottom panel of (B) are alanine check points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

from the HN(C)N-A spectrum, where as they look similar in the HNN-A spectrum. Similarly, PAA and AAP patterns are similar in HNN-A, but are distinctly different in HN(C)N-A. PAP and XAP would look similar in HN(C)N-A, but they can be readily distinguished from the HNN-A spectrum, and so on.

Thus glycines, alanines, and prolines can be used as starting points or check points during sequential walks along the polypeptide chain. Whenever, complete assignments are hampered for some reasons, by any of the methods, the fixed points will serve at least as partial monitors and enable local characterization of the protein.

The above ideas have been experimentally tested with a few proteins, and we show here the data on DLC8, a component of the motor protein, dynein [24].

Sequential walk through alanine check points: DLC8 protein has 89 residues and there are four glycines. However, it has a long stretch of 51 residues which does not have any glycines. It is here the new alanine check points become extremely useful. Fig. 4 shows identical (F_1–F_3) strips through HNN and HNN-A spectra for the monomeric protein DLC8. The changes in the patterns due to alanines can be clearly seen. Fig. 5 shows identical sequential walks through HNN (top panel) and HNN-A (bottom panel) spectra for both folded (A) and unfolded (B) DLC8. The specific patterns due to alanine are clearly evident. A summary of all the observed sequential connectivities in the HNN-A spectrum for the folded DLC8 protein is shown in Fig. 6. The glycine, proline and alanines positions along the sequence have been highlighted in the figure. Finally, Fig. 7, shows regions from the HN(C)N-A spectrum to illustrate the alanine check points, in that spectrum.

4. Conclusions

We have described here new pulse sequences which have significant advantages for rapid assignment in the context of structural genomics and protein folding studies. We have demonstrated the successful application of these new pulse sequences HNN-A and HN(C)N-A for the protein DLC8. These sequences increase the number of start/check points when compared to HNN and/or HN(C)N and hence help in pacing up the resonance assignments in proteins. However, as discussed in the original description of the HNN and HN(C)N techniques [6,1], one of the concerns is the low sensitivity of the experiments. While this is not a serious problem in unfolded proteins and smaller proteins, it may become a matter of concern for larger folded proteins which have fast relaxation rates. Deuteration may help overcome this problem by reducing the relaxation losses through the pulse sequences. Even so, the results with non-deuterated proteins presented here indicate that the experiments are still practical and reasonable sensitivities are evident. In the present case, the experiments have been performed on a VARIAN 600 MHz *Inova* NMR spectrometer with a S/N ratio of ~540, where as in the

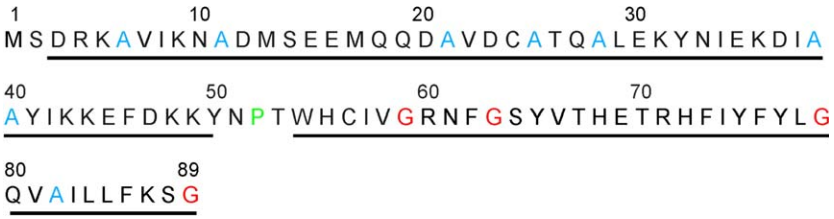

Fig. 6. A summary of all the observed sequential connectivities in the HNN-A spectrum for the folded protein shown by underlining along the sequence. The glycines are marked by red, alanines by blue and prolines by green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

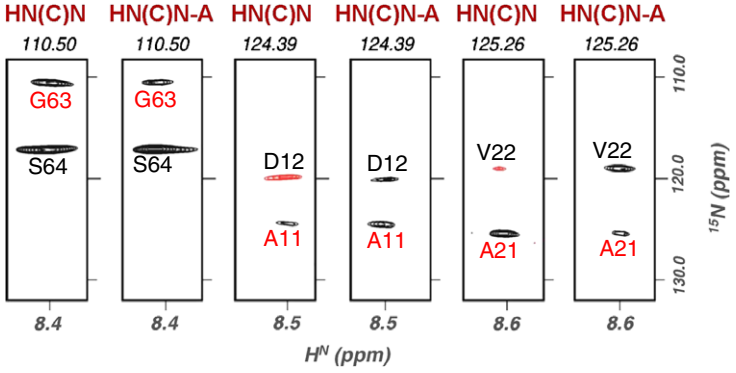

Fig. 7. Identical F_1–F_3 strips through HN(C)N and HN(C)N-A spectra of the protein DLC8 at pH3, 27 °C. The numbers on top of each panel are the F_2 chemical shifts which help to identify the diagonal ($F_1 = F_2$) peaks. Black contours are positive peaks and red contours are negative peaks. The residue corresponding to the diagonal in each strip has been indicated by red. The distinctive behavior of alanines is evident for the residues A11 and A21 as they show negative diagonal (i) a positive cross ($i + 1$) peak in HN(C)N-A.

modern NMR spectrometers ratios of 1300 are easily available and with the availability of 'cryoprobes,' S/N ratios of as much as 6000 have been realized on 600 MHz spectrometers. At higher magnetic field strengths of up to 900 MHz (in proton frequency units), as are currently available, the sensitivities are even higher. Thus, sensitivity of the HNN-A and HN(C)N-A techniques is not as much of an issue and experiments can be successfully carried out on large proteins even without deuteration.

References

[1] S.C. Panchal, N.S. Bhavesh, R.V. Hosur, Improved 3D triple resonance experiments, HNN and HN(C)N, for H^N and ^{15}N sequential correlations in (^{13}C, ^{15}N) labeled proteins: application to unfolded proteins, J. Biomol. NMR 20 (2001) 135–147.

[2] R. Weisemann, H. Ruterjans, W. Bermel, 3D triple-resonance NMR techniques for the sequential assignment of NH and ^{15}N resonances in ^{15}N- and ^{13}C-labelled proteins, J. Biomol. NMR 3 (1993) 113–120.

[3] S. Grzesiek, J. Anglister, H. Ren, A. Bax, Carbon-13 line narrowing by deuterium decoupling in deuterium/carbon-13/nitrogen-15 enriched proteins. Application to triple resonance 4D J connectivity of sequential amides, J. Am. Chem. Soc. 115 (1993) 4369–4370.

[4] H. Matsuo, E. Kupce, H. Li, G. Wagner, Use of selective C alpha pulses for improvement of HN(CA)CO-D and HN(COCA)NH-D experiments, J. Magn. Reson. B 111 (1996) 194–198.

[5] C. Bracken, A.G. Palmer III, J. Cavanagh, (H)N(COCA)NH and HN(COCA)NH experiments for ^1H–^{15}N backbone assignments in $^{13}C/^{15}N$-labeled proteins, J. Biomol. NMR 9 (1997) 94–100.

[6] N.S. Bhavesh, S.C. Panchal, R.V. Hosur, An efficient high-throughput resonance assignment procedure for structural genomics and protein folding research by NMR, Biochemistry 40 (2001) 14727–14735.

[7] A. Chatterjee, N.S. Bhavesh, S.C. Panchal, R.V. Hosur, A novel protocol based on HN(C)N for rapid resonance assignment in (^{15}N, ^{13}C) labeled proteins: implications to structural genomics, Biochem. Biophys. Res. Commun. 293 (2002) 427–432.

[8] M. Bozzi, M. Bianchi, F. Sciandra, M. Paci, B. Giardina, A. Brancaccio, D.O. Cicero, Structural characterization by NMR of the natively unfolded extracellular domain of beta-dystroglycan: toward the identification of the binding epitope for alpha-dystroglycan, Biochemistry 42 (2003) 13717–13724.

[9] A. Chatterjee, P. Mridula, R.K. Mishra, R. Mittal, R.V. Hosur, Folding regulates autoprocessing of HIV-1 protease precursor, J. Biol. Chem. 280 (2005) 11369–11378.

[10] J. Juneja, N.S. Bhavesh, J.B. Udgaonkar, R.V. Hosur, NMR identification and characterization of the flexible regions in the 160 kDa molten globule-like aggregate of barstar at low pH, Biochemistry 41 (2002) 9885–9899.

[11] C.J. Macdonald, K. Tozawa, E.S. Collins, C.N. Penfold, R. James, C. Kleanthous, N.J. Clayden, G.R. Moore, Characterisation of a mobile protein-binding epitope in the translocation domain of colicin E9, J. Biomol. NMR 30 (2004) 81–96.

[12] S.C. Panchal, D.A. Kaiser, E. Torres, T.D. Pollard, M.K. Rosen, A conserved amphipathic helix in WASP/Scar proteins is essential for activation of Arp2/3 complex, Nat. Struct. Biol. 10 (2003) 591–598.

[13] P. Teriete, S. Banerji, M. Noble, C.D. Blundell, A.J. Wright, A.R. Pickford, E. Lowe, D.J. Mahoney, M.I. Tammi, J.D. Kahmann, I.D. Campbell, A.J. Day, D.G. Jackson, Structure of the regulatory hyaluronan binding domain in the inflammatory leukocyte homing receptor CD44, Mol. Cell 13 (2004) 483–496.

[14] K. Tozawa, C.J. Macdonald, C.N. Penfold, R. James, C. Kleanthous, N.J. Clayden, G.R. Moore, Clusters in an intrinsically disordered protein create a protein-binding site: the TolB-binding region of colicin E9, Biochemistry 44 (2005) 11496–11507.

[15] X. Zhang, Y. Xu, J. Zhang, J. Wu, Y. Shi, Structural and dynamic characterization of the acid-unfolded state of hUBF HMG box 1 provides clues for the early events in protein folding, Biochemistry 44 (2005) 8117–8125.

[16] J. Wirmer, H. Schwalbe, Angular dependence of $^1J_{N_iC_i^\alpha}$ and $^2J_{N_iC_{i-1}^\alpha}$ coupling constants measured in J-modulated HSQCs, J. Biomol. NMR 23 (2002) 47–55.

[17] N.S. Bhavesh, A. Chatterjee, S.C. Panchal, R.V. Hosur, Application of HN(C)N to rapid estimation of $^1J_{NC^\alpha}$ coupling constants correlated to ψ torsion angles in proteins: implication to structural genomics, Biochem. Biophys. Res. Commun. 311 (2003) 678–684.

[18] J. Star-Lack, S.J. Nelson, J. Kurhanewicz, L.R. Huang, D.B. Vigneron, Improved water and lipid suppression for 3D PRESS CSI using RF band selective inversion with gradient dephasing (BASING), Magn. Reson. Med. 38 (1997) 311–321.

[19] D. Rosenfeld, S.L. Panfil, Y. Zur, Design of adiabatic pulses for fat-suppression using analytic solutions of the Bloch equation, Magn. Reson. Med. 37 (1997) 793–801.

[20] R. Crouch, R.D. Boyer, R. Johnson, K. Krishnamurthy, Broadband and band-selective IMPRESS-gHMBC: compensation of refocusing inefficiency with synchronized inversion sweep, Magn. Reson. Chem. 42 (2004) 301–307.

[21] S.A. Bradley, H. Hu, K. Krishnamurthy, C.E. Hadden, A band-selective composite gradient: application to DQF-COSY, J. Magn. Reson. 174 (2005) 110–115.

[22] M.S. Silver, R.I. Joseph, D.I. Hoult, Selective spin inversion in nuclear magnetic resonance and coherent optics through an exact solution of the Bloch–Riccati equation, Phys. Rev. A 31 (1985) 2753–2755.

[23] W. Peti, L.J. Smith, C. Redfield, H. Schwalbe, Chemical shifts in denatured proteins: resonance assignments for denatured ubiquitin and comparisons with other denatured proteins, J. Biomol. NMR 19 (2001) 153–165.

[24] E.L. Holzbaur, R.B. Vallee, DYNEINS: molecular structure and cellular function, Annu. Rev. Cell Biol. 10 (1994) 339–372.

[25] A.J. Shaka, Composite pulses for ultra broadband inversion, Chem. Phys. Lett. 120 (1985) 201–205.